

# Robust Wiener Filtering for Adaptive Optics

*Lisa A. Poyneer*

*LLNL Technical Document  
June 18, 2004*

*U.S. Department of Energy*

Lawrence  
Livermore  
National  
Laboratory

## **DISCLAIMER**

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

This is a preprint of a paper intended for publication in a journal or proceedings. Since changes may be made before publication, this preprint is made available with the understanding that it will not be cited or reproduced without the permission of the author.

# Robust Wiener Filtering for Adaptive Optics

Lisa A. Poyneer

## 1. Adaptive Optics Signal Processing

In many applications of optical systems, the observed field in the pupil plane has a non-uniform phase component. This deviation of the phase of the field from uniform is called a phase aberration. In imaging systems this aberration will degrade the quality of the images. In the case of a large astronomical telescope, random fluctuations in the atmosphere lead to significant distortion. These time-varying distortions can be corrected using an Adaptive Optics (AO) system, which is a real-time control system composed of optical, mechanical and computational parts. Adaptive optics is also applicable to problems in vision science, laser propagation and communication. For a high-level overview, consult this web site.<sup>1</sup> For an in-depth treatment of the astronomical case, consult these books.<sup>2,3</sup>

Our experimental data is from Lick Observatory on Mt. Hamilton, near San Jose, CA. The 120-inch (3-meter) Shane telescope has an AO system which was developed by LLNL and is jointly maintained with the observatory. In the Lick AO system the phase aberration is measured using a wave-front sensor (WFS). The sensor divides the pupil of the telescope into smaller sections called subapertures, by a small array of square lenslets. Each lenslet forms a spot image of the reference source on a small region of a CCD. As the phase aberration of the field in the subaperture changes, the spot moves. This motion is directly related to the spatial slope of the phase across the subaperture. By estimating this movement of the spot in both directions on the CCD, the x- and y-slopes of the wavefront in the subaperture can be estimated. These slopes are then sent to a reconstruction algorithm, which generates the phase across the entire pupil based on the local slopes. This reconstructed phase is then applied to a flexible optic which conjugates the phase aberration of the field. This process is done in closed loop and at rates of up to 500 Hz. Full telemetry of the WFS slope estimates is available in bursts of 4096 consecutive time steps.

For this final project, we will focus in on a single area of the AO system: the WFS measurements. In this case the slope telemetry data provides a noisy measurement of the phase aberration which the system is either observing in open-loop or correcting in closed-loop. This measurement of phase slope through time and across the pupil is important for many areas of my research. For this project I will not discuss what I will do with the estimated slopes for my research. I will instead focus on the best way to filter these noisy measurements using filtering techniques as covered in EEC 263 and the literature for this final project.

## 2. Robust Wiener Filtering

Our standard model for signal estimation is that we observe the linear combination of two real-valued, zero-mean, uncorrelated wide-sense stationary random processes being the signal  $s(t)$  and the noise  $n(t)$ . The power spectral densities (PSDs) of both of these processes exist and are known as  $S(\omega)$  and  $N(\omega)$ . We observe the noisy signal as  $y(t) = s(t) + n(t)$ . In this case, the optimal

linear filter to estimate  $s(t)$  from  $y(t)$  is the Wiener filter. In the IIR case (which is the only one we consider in this project) the optimal linear filter  $h(t)$  such that

$$\hat{s}(t) = \int_{-\infty}^{\infty} y(\tau)h(t - \tau)d\tau, \quad (1)$$

is specified to have a frequency response  $H(\omega)$  related to the PSDs of the random processes as

$$H(\omega) = \frac{S(\omega)}{S(\omega) + N(\omega)}. \quad (2)$$

This is somewhat an abuse of notation since  $H(\omega)$  is the Fourier Transform of  $h(t)$  and  $S(\omega)$  is the Fourier transform of the autocovariance function of the random process  $s(t)$ , but it is consistent with the literature. See the book by Hayes<sup>4</sup> for a thorough treatment for the theoretical basis of Wiener filtering. In most real situations the exact PSDs of the signal and noise are not known, hence the optimum filter cannot be designed.

The robust filtering approach uses the framework of a class of possible signal and noise PSDs. An exact PSD is not available, but may be known to be a member of some defined class. We define the general classes as  $\mathcal{S}$  for the signal and  $\mathcal{N}$  for the noise. Given the classes, and some constraints upon them, a robust estimation filter can be derived. The robust estimator is a filter that is not optimum for the nominal case, but provides acceptable performance for the non-nominal case. For Wiener filtering this approach is summarized in a tutorial paper<sup>5</sup> by Kassam and Poor. In-depth examination of specific types of classes occurs in papers by Kassam and Lim<sup>6</sup> and Poor.<sup>7</sup> In all cases, the framework used is that of *minimax*. The filter is chosen such that its maximum error for any signal in the class is the minimum over all filters. The robust filter has the best worst-case performance of all possible filters.

Note that though equations presented in this report (and in the literature) assume the most general case of continuous random processes of infinite extent, my actual simulations and data analysis are on finite-length sampled data. In this case I am post-processing sampled data of length 4096 samples. Because I am not faced with computational constraints, I will use the IIR Wiener filter under the assumption that my sampled data series are periodic every 4096 samples. This allows the optimal filters to be directly obtained from the estimated PSDs, which are in turn obtained with the Discrete Fourier Transform. The filters are applied in the frequency domain to the Fourier transform of the input signals.

## A. PSD classes

In order to get the robust filter, it is necessary to define the classes from which the signal and noise PSDs can be drawn. Given the WFS data which is observed, I have chosen a single type of model, with two variations. I used the bounded class model (or band model), which is, as Kassam and Lim note, reasonable for the case of PSDs estimated from real data. It is also reasonable given the underlying physics, which will be discussed below.

In this model, upper and lower bounds are placed on the signal and noise PSDs. For the signal, upper and lower bounds on the PSD are specified such that

$$\mathcal{S} \equiv \{S(\omega) | S_L(\omega) \leq S(\omega) \leq S_U(\omega)\}, \quad (3)$$

where  $S_L(\omega)$  is the lower bound and  $S_U(\omega)$  is the upper bound. For Kassam and Poor, a second constraint is added on the total power in any member of the class such

$$\mathcal{S}_P \equiv \left\{ S(\omega) \left| S_L(\omega) \leq S(\omega) \leq S_U(\omega) \text{ and } \frac{1}{2\pi} \int_{-\pi}^{\pi} S(\omega) d\omega = \sigma_s^2 \right. \right\}, \quad (4)$$

where  $\sigma_s^2$  is specified. The noise classes  $\mathcal{N}$  and  $\mathcal{N}_P$  are likewise specified, with the bounds  $N_L(\omega)$  and  $N_U(\omega)$  and power  $\sigma_n^2$ .

In the more general case discussed by Poor, the only constraint on the class is that it is convex. A convex set can be defined in the general case as a subspace where the line connecting any two points is also in that subspace. Or more directly, for all  $x, y \in C$ , where  $C$  is the class, for all  $t \in [0, 1]$ ,  $tx + (1 - t)y \in C$ . The bounded class model is convex, and the proof is short. At any frequency  $\omega$  a member of the bounded class is bounded from above and below as  $L \leq x \leq U$ . Since  $t \in [0, 1]$  we obtain

$$tL + (1 - t)L \leq tx + (1 - t)y \leq tU + (1 - t)U, \quad (5)$$

which becomes

$$L \leq tx + (1 - t)y \leq U, \quad (6)$$

and the convexity is shown.

A specific model for the bounded classes  $\mathcal{S}$  and  $\mathcal{N}$  is needed to derive the robust filters. This model is based on both real observations and theoretical analysis of atmospheric phase distortions. A good model for the temporal power spectrum of WFS measurements of the atmosphere is that the true slope measurements follow a temporal power law. For an in-depth discussion of the underlying physics, see Chapter 3 of Hardy.<sup>2</sup> Depending on whether low-order correction is applied in the system, very low temporal frequencies may fall short of the true power law. This power law is parameterized by a constant  $r_0$  which describes the power of the turbulence. This parameter is known to fluctuate on all time scales in the atmosphere. In addition, the true PSD may deviate somewhat from the power law. The noise on the measurements arises from noise on the CCD and is nominally to be temporally uncorrelated. The portion due to photon noise will vary with time as the number of received photons varies. The read noise portion due to the CCD itself will remain relatively constant. A bounded model was constructed for both the signal and noise PSDs. For these simulations, the nominal signal PSD for uncorrected atmospheric turbulence is

$$S_N(\omega) = r_0(\omega^2)^{-11/12}, \quad (7)$$

where  $r_0$  is set to a baseline total power of  $2\pi\sigma_s^2$ . The upper and lower bounds are set as a bandwidth around the nominal set by the parameter  $b \in [0, 1]$ , such that

$$S_L(\omega) = (1 - b)S_N(\omega), \quad S_U(\omega) = (1 + b)S_N(\omega). \quad (8)$$

The noise class is likewise defined with the same parameter  $b$ , where the nominal noise PSD is

$$N_N(\omega) = C, \quad (9)$$

with a baseline total power of  $2\pi\sigma_n^2$ . It is also assumed that the classes are distributed such that over a large random sampling for a specific  $b$ , the mean PSDs will approach the nominal PSDs

$S_B(\omega)$  and  $N_B(\omega)$ . This allows the use of a nominal filter derived from the sample mean (which is what would be done with experimental data) as a comparison filter to the robust filter.

For the power-constrained classes  $\mathcal{S}_P$  and  $\mathcal{N}_P$ , any member can deviate from the nominal, but deviations that produce more power at a specific frequency must be compensated for by deviations at other frequencies with lower power. Since this power conservation is not necessarily reasonable given the behavior of the atmosphere, the non-power constrained class  $\mathcal{S}$  and  $\mathcal{N}$  will also be considered. In the following subsections the robust filters will be derived. In the power constrained case, the results of Kassam and Lim allow a direct calculation of the robust filter given their formulae. In the non-power constrained case, the general approach of Poor is applied to our specific case.

## B. Robust filter for the power constrained model

Kassam and Lim derived the exact robust filter for the power-constrained bounded class model. This section summarizes their results and applies them to our model. A discussion of the behavior of the robust filter follows.

The first step in deriving the robust filter is to find the least-favorable PSDs  $S_R(\omega)$  and  $N_R(\omega)$ . The robust filter for the whole class is the optimal Wiener filter for this specific case. The filters are formed by examining the ratios of the signal and noise PSDs and solving for the required total power. The least-favorable PSDs are found by partitioning the frequency axis into specific sets based on ratios of PSDs. Kassam and Lim present four different decompositions depending on the structure of the signal and noise PSD class. For both our observational data and the simulation model for a range of SNRs, the PSD class will fall into their case (b) of Theorem 1 (see p. 176 of their paper).

First the frequency axis is partitioned into four disjoint sets based the PSDs and the value of a parameter  $k$ . This parameter  $k$  is compared to the three different ratios of the signal power to the noise power. These three signal-to-noise metrics (SNRs) are the best possible SNR  $S_U(\omega)/N_L(\omega)$ , the worst possible SNR  $S_L(\omega)/N_U(\omega)$  and the lower SNR  $S_L(\omega)/N_L(\omega)$ . In particular, the sets, which are a function of the parameter  $k$  are defined as

$$\beta_2(k) \equiv \left\{ \omega \mid k \leq \frac{S_L(\omega)}{N_U(\omega)} \right\}, \quad (10)$$

$$\beta_1(k) \equiv \left\{ \omega \mid \frac{S_L(\omega)}{N_U(\omega)} < k \leq \frac{S_L(\omega)}{N_L(\omega)} \right\}, \quad (11)$$

$$\alpha_1(k) \equiv \left\{ \omega \mid \frac{S_L(\omega)}{N_L(\omega)} < k \leq \frac{S_U(\omega)}{N_L(\omega)} \right\}, \quad (12)$$

$$\alpha_2(k) \equiv \left\{ \omega \mid \frac{S_U(\omega)}{N_L(\omega)} < k \right\}. \quad (13)$$

These four sets completely cover the frequency axis. A sample decomposition (for a case similar to that of our model) is shown in Fig. 1.

The value of  $k$  is found by solving the following equation, which integrates the PSDs over these segments.

$$k = \frac{2\pi\sigma_s^2 - \int_{\beta_2(k)} S_L(\omega)d\omega - \int_{\alpha_2(k)} S_U(\omega)d\omega}{2\pi\sigma_n^2 - \int_{\beta_2(k)} N_U(\omega)d\omega - \int_{\alpha_2(k)} N_L(\omega)d\omega}. \quad (14)$$

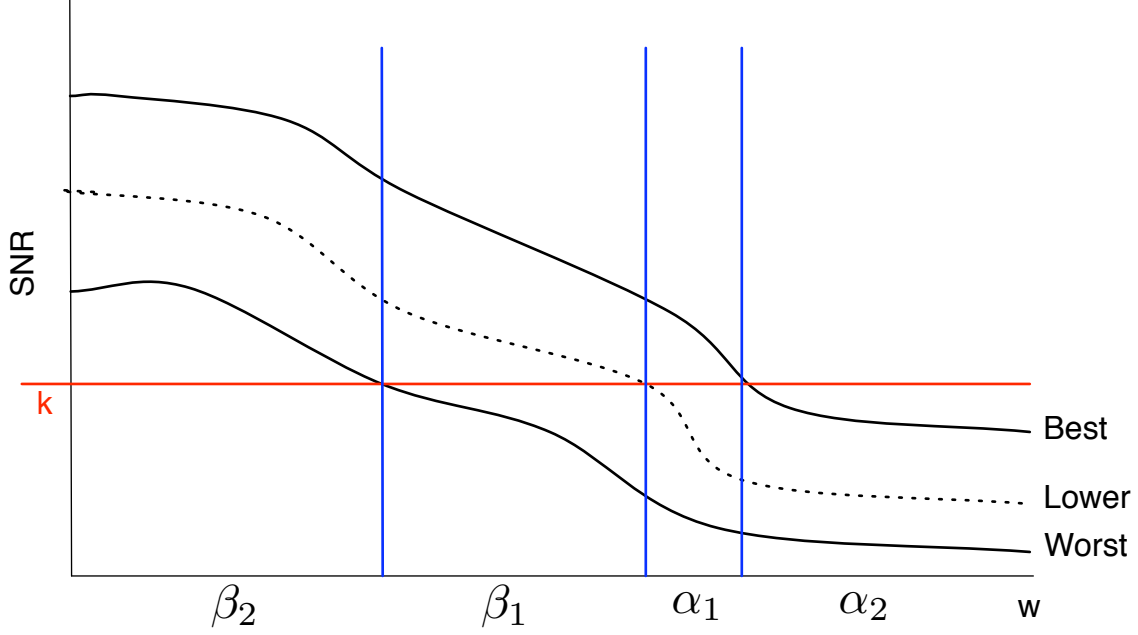


Fig. 1. For a given  $k$ , shown by the red line, the frequency axis is partitioned into four sets, based on the three SNR metrics.

This equation balances out the power in the different regions.

Once this  $k$  is found, the least-favorable PSDs are given by the following equations, where the PSDs  $S_e(\omega)$  and  $N_e(\omega)$  and non-negative and satisfy  $S_e(\omega) = kN_e(\omega)$ ,

$$S_R(\omega) = \begin{cases} S_L(\omega) & \omega \in \beta_2(k) \\ S_L(\omega) + S_e(\omega) & \omega \in \beta_1(k) \\ kN_L(\omega) + S_e(\omega) & \omega \in \alpha_1(k) \\ S_U(\omega) & \omega \in \alpha_2(k) \end{cases}, \quad (15)$$

$$N_R(\omega) = \begin{cases} N_U(\omega) & \omega \in \beta_2(k) \\ \frac{1}{k}S_L(\omega) + N_e(\omega) & \omega \in \beta_1(k) \\ N_L(\omega) + N_e(\omega) & \omega \in \alpha_1(k) \\ N_L(\omega) & \omega \in \alpha_2(k) \end{cases}. \quad (16)$$

Now the robust filter is simply the Wiener filter for the least-favorable PSDs,

$$H_R(\omega) = \frac{S_R(\omega)}{S_R(\omega) + N_R(\omega)}. \quad (17)$$

Based on the specifications given above for the least-favorable PSDs, the filter is simply

$$H_R(\omega) = \begin{cases} \frac{S_L(\omega)}{S_L(\omega) + N_U(\omega)} & \omega \in \beta_2(k) \\ \frac{k}{k+1} & \omega \in \beta_1(k), \omega \in \alpha_1(k) \\ \frac{S_U(\omega)}{S_U(\omega) + N_L(\omega)} & \omega \in \alpha_2(k) \end{cases}. \quad (18)$$

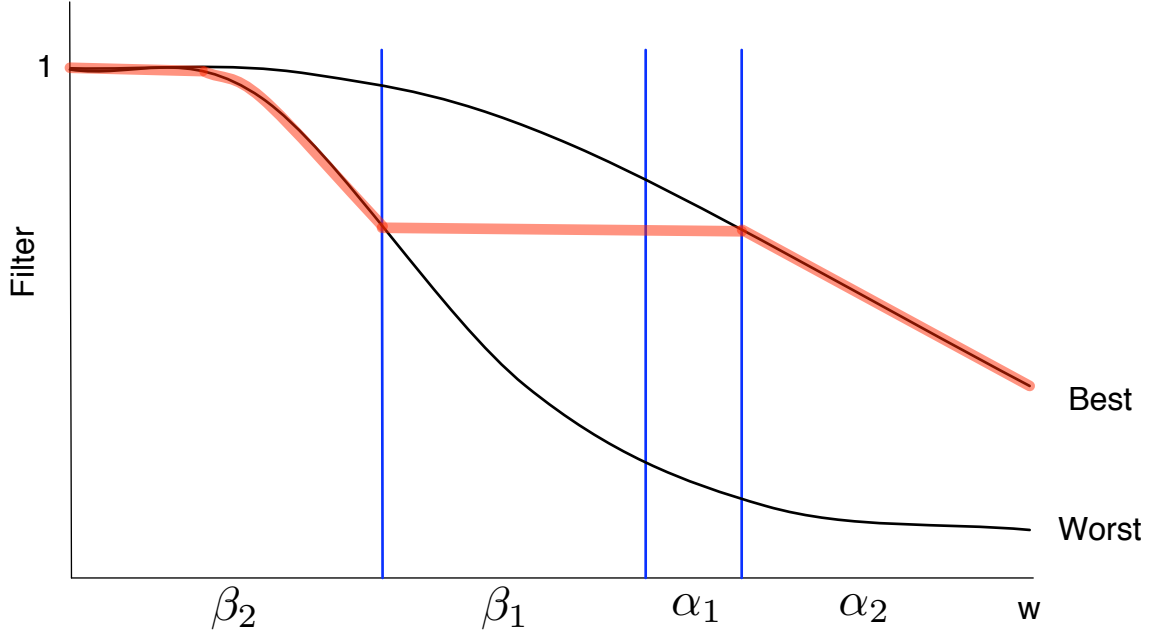


Fig. 2. The robust filter (red line) is composed of different parts. It is the optimal Wiener filter to the least-favorable PSDs. For  $\beta_2(k)$ , it follows the Wiener filter for the worst SNR. For  $\alpha_2(k)$  it follows the Wiener filter for the best SNR. In the middle it has a constant value connecting the two.

This means that where the signal is strongest,  $\omega \in \beta_2(k)$ , the filter is pessimistically assigned to the lowest possible SNR scenario. In the middle,  $\omega \in \beta_1(k), \omega \in \alpha_1(k)$  the filter is constant, set by the  $k$  value. In the noise band,  $\omega \in \alpha_2(k)$ , the filter is set to the best SNR. This is illustrated in Fig. 2. This robust filter corresponds to a conservative approach to the SNRs at various frequencies. The optimal Wiener filter can be written as

$$H(\omega) = \frac{1}{1 + 1/\text{SNR}}. \quad (19)$$

When the SNR is high, the filter gain approaches one. When the SNR is low, the gain approaches zero. In this high region, the robust filter guards against the possibility that the SNR is lower than expected by lowering the gain. In the low SNR region, the robust filter guards against the possibility that the SNR is higher than expected by raising the gain.

### C. Robust filter for the non-power constrained model

In this more general case the minimax problem has not been solved explicitly. The framework of Poor is necessary and solve the problem (either numerically or analytically) for our specific case.

Given the convex classes  $\mathcal{S}$  and  $\mathcal{N}$ , the least-favorable spectral pair  $(S_R(\omega), N_R(\omega))$  is defined as the pair for which the error for that pair when optimally filtered is the maximum over all pairs from the entire class. Then the robust filter is simply the optimal Wiener filter  $H_R(\omega)$  for this pair. (Kassam and Lim solve the specific case of this as described above.)



In this case, since the classes are convex, the maximization can be solved. However, to make this tractable, we will further constrain our model to make the members of the class  $\mathcal{S}$  satisfy the condition that given nominal PSD  $S_N(\omega)$ , any member of the class is

$$kS_N(\omega), k \in [1 - b, 1 + b], \quad (20)$$

and for  $\mathcal{N}$ , with nominal PSD  $N_N(\omega)$ ,

$$lN_N(\omega), l \in [1 - b, 1 + b]. \quad (21)$$

We call this new class pair  $(\mathcal{S}_B, \mathcal{N}_B)$ . This model assumes that the temporal structure of the signal and noise PSDs remains constant, but the total power can fluctuate. Given this model, finding the least-favorable PSDs reduces to solving maximization

$$\max_{k,l} \int_{-\infty}^{\infty} \frac{klS_N(\omega)N_N(\omega)}{kS_N(\omega) + lN_N(\omega)} d\omega. \quad (22)$$

This maximization is derived from the formula for the MSE of Wiener filtering.

Solving this maximization problem leads to the solution that  $S_R(\omega) = (1 + b)S_N(\omega)$  and  $N_R(\omega) = (1 + b)N_N(\omega)$ . This makes the robust filter be the same as the Wiener filter for the nominal spectra

$$H_R(\omega) = \frac{S_N(\omega)}{S_N(\omega) + N_N(\omega)}. \quad (23)$$

This result is quite different from the previous case with the power constrained. In this case the robust filter is actually the nominal Wiener filter.

#### D. Summary

The bounded spectral class model  $(\mathcal{S}, \mathcal{N})$  is used for our signal estimation problem. It assumes that the nominal signal PSD is a power law in temporal frequency and that the nominal noise PSD is a constant. The upper and lower bounds on the classes are defined by a bandwidth parameter  $b \in [0, 1)$  such that the upper bound is  $(1 + b)$  times the nominal and the lower bound is  $(1 - b)$  times the nominal. For the power-constrained case of Kassam and Poor  $(\mathcal{S}_P, \mathcal{N}_P)$ , any member of  $\mathcal{S}$  has the same total power as the nominal, and likewise for the noise class. Using the results of Kassam and Lim, the robust filter can be found. In our case it has the structure that at low frequencies (where the SNR is highest) it follows the Wiener filter for the worst-possible SNR case. At mid-range frequencies (and middle SNRs) the filter is constant. At high frequencies (where the SNR is lowest) the filter follows the best possible SNR response. This results in a filter that is conservative in both bands.

If the power constraint is disregarded and instead we assume that any member of  $\mathcal{S}$  and  $\mathcal{N}$  is simply a scaling of the nominal PSD (limited of course by the upper and lower bounds), we can derive the robust filter following Poor. In this case  $(\mathcal{S}_B, \mathcal{N}_B)$  the robust filter is exactly the Wiener filter for the nominal PSDs.

In reference to experimental data, there are now two possible approaches. It is first necessary to estimate the nominal PSD and its bounds from the data. Given these bounds, we can assume

$(\mathcal{S}_P, \mathcal{N}_P)$  and find the robust filter given the upper and lower bounds on the PSDs. Or we can assume  $(\mathcal{S}_B, \mathcal{N}_B)$  and find the robust filter from the mean PSDs. In the next section we study the behavior of these two different robust filters given the signal classes to address the question of which approach might be most useful for our experimental data.

### 3. Simulation Results

Simulation is used to explore the performance of the robust filters for measurements with spectral uncertainty, as described above. This section addresses the comparative performance of the different robust filters for three PSD models. This performance depends on the model distributions. The impact of the estimation of PSDs from data to obtain the filters is also studied.

The majority of simulations were done at an SNR of 10 with the sum of  $S(\omega) = 500$  and the sum of  $N(\omega) = 50$  (which is close to the Lick experimental data). Some experiments were repeated at SNR of 2 with  $S(\omega) = 100$ . Two different distributions for the power-constrained classes  $(\mathcal{S}_P, \mathcal{N}_P)$  and one for  $(\mathcal{S}_B, \mathcal{N}_B)$  are used in simulations. For  $(\mathcal{S}_P, \mathcal{N}_P)$  the realizations of the PSDs were created with the following model

$$S_i(\omega) = (1 + k(\omega))S_N(\omega), \quad k(\omega) \sim \text{IID Uniform}[-b, b], \quad (24)$$

and

$$N_i(\omega) = (1 + l(\omega))N_N(\omega), \quad l(\omega) \sim \text{IID Uniform}[-b, b], \quad (25)$$

then adjusted slightly by a multiplicative factor to meet the power constraint. (The random numbers were created for positive  $\omega$  and reflected to preserve the symmetry of the PSDs.) The second version had  $k(\omega)$  and  $l(\omega)$  as scaled cosines with a random frequency and phase shift,

$$S_i(\omega) = (1 + k(\omega))S_N(\omega), \quad (26)$$

$$k(\omega) = b \cos(2\pi\omega/T_1) T_1 \sim \text{IID Uniform}[100, 500], \quad (27)$$

and

$$N_i(\omega) = (1 + l(\omega))N_N(\omega), \quad (28)$$

$$l(\omega) = b \cos(2\pi\omega/T_2 + \phi), \quad T_2 \sim \text{IID Uniform}[200, 1000], \quad \phi \sim \text{IID Uniform}[0, 2\pi], \quad (29)$$

again with slight modifications to ensure the power constraint. These two models represent different ways the PSDs can vary from the nominal. In the first case, each frequency varies independently. In the second case the variations at different frequencies are correlated. These are primarily useful as a way to get a useful subset of the total class. For the nonconstrained case the model as specified above (in Eq. 20) was directly implemented, with the multiplicative constant uniformly distributed between  $[1 - b, 1 + b]$ .

In the following results the white-noise distributions are referred to as *Random*. The sin-based distributions are *Sin* and the non-power constrained are ‘Noncon’. The robust filter for  $(\mathcal{S}_P, \mathcal{N}_P)$  is called *Robust*, the robust filter for  $(\mathcal{S}_B, \mathcal{N}_B)$  (which is of course the nominal Wiener filter) is called *Nominal* and the actual Wiener filter for the specific PSD realizations is called *TrueOpt*.

### A. MSE space exploration

Given the spectral classes and the robust filters, we can calculate the MSE of estimation for a given spectral pair and an arbitrary filter. This MSE is

$$\text{MSE} = \frac{1}{2\pi} \int S_a(\omega) |1 - H_a(\omega)|^2 + N_a(\omega) |H_a(\omega)|^2 d\omega, \quad (30)$$

for any signal and noise PSDs  $(S_a(\omega), N_a(\omega))$  and filter  $H(\omega)$ . This is calculated numerically as the sum of the sampled PSDs.

Given these spectral models, the *Robust* and *Nominal* filters can be calculated once. Then 10,000 realizations from the spectral distributions are created. For each realization pair, the *TrueOpt* filter is determined. Then the MSE of estimation for each of the three filters is calculated. These results are then saved for each trial.

This is a simple Monte Carlo approach to examining the performance of the various filtering options for the spectral models. In each case a large number of realizations is used to adequately sample the class space. For each realization the true optimal MSE of estimation is compared to what would be obtained if the either of the two robust filters were used. Note that the *Robust* is derived only for the power-constrained model, but is applied to both models, and *Nominal* is derived for the unconstrained model only but also applied to both.

These trials were conducted for  $b = .25$  and  $b = .5$ , both at the SNR of 10 in terms of total power, as specified above. For both the *Random* and *Sin* distributions of the power-constrained model  $(\mathcal{S}_P, \mathcal{N}_P)$ , the *Nominal* filter significantly out-performed the *Robust* filter. I verified that in the specific case of the least-favorable PSDs, *Robust* does outperform *Nominal*, but it should, since it is the Wiener filter solution in this case. Given these probability distributions for the classes, it is clear that the set of PSD pairs which are unfavorable is actually very small. One of the assumptions that motivates robust filtering is that performance for the nominal filter falls off rapidly as the PSDs diverge from the nominal (see Kassam and Poor, for example). This does not appear to be the case here. Plots comparing the MSEs are shown in Figs. 3 and 4 for the  $b = .5$  case. In the non-power-constrained model  $(\mathcal{S}_B, \mathcal{N}_B)$  *Robust* did well for a small portion of phase space: it performed better than *Nominal* about 10% of the time. But for most cases it performed worse. This is an allowable result. The robust framework only says that for PSDs other than the least-favorable, the robust filter (which in this case is *Nominal*) does well, but other filters may do better.

So under all three distributions, the best filter to use is the *Nominal* filter. These simulations were repeated for an SNR of 2, and the results were similar. This, however, assumes that the nominal, lower and upper bound PSDs are all known for both the signal and the noise. In real data processing these PSDs are not known but are estimated from the data itself. A second useful simulation set to do is to estimate the PSDs from a small set of observations.

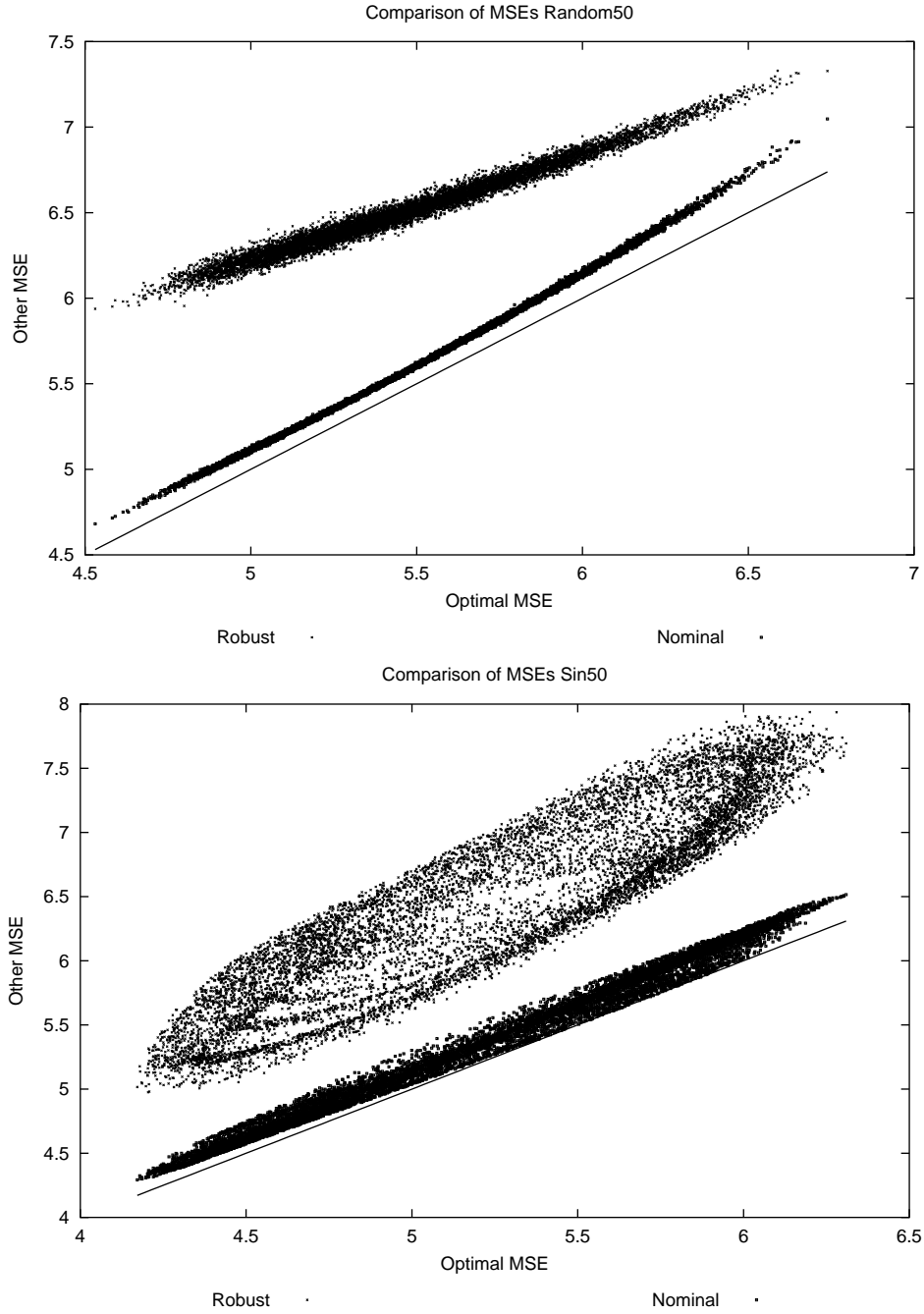


Fig. 3. For thousands of random PSD pairs from the power-constrained models *Random* and *Sin*, the optimal MSE with Wiener filtering is compared to the MSE for the *Robust* filter and the *Nominal* filter. In all cases the *Nominal* filter has lower MSE than the *Robust* filter.

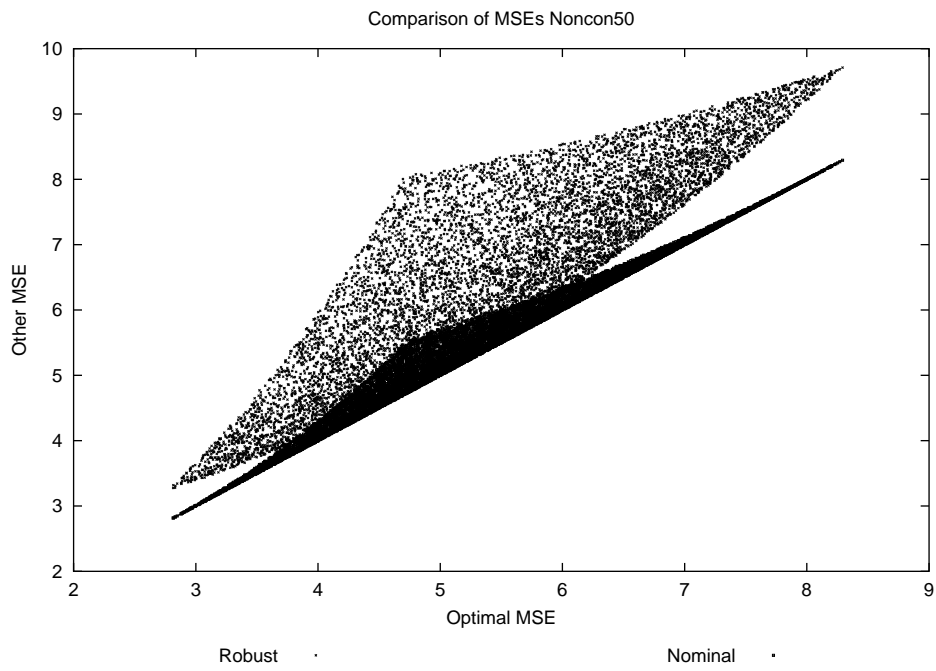


Fig. 4. For thousands of random PSD pairs from the non-power-constrained model ‘Noncon’, the optimal MSE with Wiener filtering is compared to the MSE for the *Robust* filter and the *Nominal* filter. In most cases the *Nominal* filter has lower MSE than the *Robust* filter.

## B. Performance with estimated PSDs

For this simulation, which examines the affects of having to estimate the PSDs from observational data, a similar approach was used to the above. In this case a small number of random PSDs pairs were generated. For each pair, a single random noisy signal realization was produced. Using these signals, the PSDs were estimated.

First the overall PSD of signal plus noise was estimated using the averaged periodogram method, which is the standard method of averaging the magnitude-squared of the DFTs of all the signals. Then the constant noise PSD was fit to the mean of the high-frequency periodogram. The lower and upper noise bounds were fit to the same region as the mean PSD minus or plus the sample standard deviation of the periodogram at those frequencies. Using the uncorrelatedness of the two PSDs, the noise PSDs were then subtracted from the total mean, lower and upper PSDs. Absolute value was taken to guard against negative results. No smoothing was done, as this tends to corrupt the lowest-frequency results due to the power law.

This PSD estimation procedure of course has uncertainty and possibly systematic errors. The first error is that the true nominal signal PSD is usually over-estimated at the highest frequencies where the noise predominates. Second, finding the lower and upper bounds is somewhat problematic. Each individual periodogram is a noisy estimate of the true PSD (for that individual case) and it deviates from the nominal PSD by both inherent variation in the PSD (due to the model) and the signal realization noise. This makes the lower and upper bound estimates more uncertain.

It is worth noting that this method of spectral estimation for the nominal PSD is reasonable, though perhaps given some more information about the true PSDs a parametric method might give less-noisy estimates. We have not implemented such an estimation technique at this time. There is most likely a better method than the basic approach for the lower and bound described above. However, since the *Robust* filter in almost all cases performed poorly, estimating these bounds is not a priority.

These PSD estimates are then used in the Monte Carlo simulation. Given the spectral estimates from the data, the *Robust* and *Nominal* filters are derived for the data set. Now each spectral pair is re-visited and the MSE of estimation is calculated for the true PSDs (again termed ‘OptTrue’) and the data-based *Robust* and *Nominal* filters. As done in the previous subsection, the MSEs are compared. These results for an SNR of 10 and  $b = .5$  are shown in Figs. 5 and 6. For only 50 random samples from the PSD classes, *Robust* filter performance is quite poor and *Nominal* also suffers. Much better spectral estimates are obtained for the 500 random samples case. In comparison to the Monte Carlo simulations where the exact robust filters were known, the *Robust* estimated from data does slightly worse than the true *Robust* filter in the 500 case. The *Nominal* filter from 500 samples is at the same level of MSE as the true *Nominal* filter. Given the extra uncertainty in the spectral estimates for the lower and upper bounds, the fact the the *Robust* filter performs less well in comparison to the ideal is to be expected.

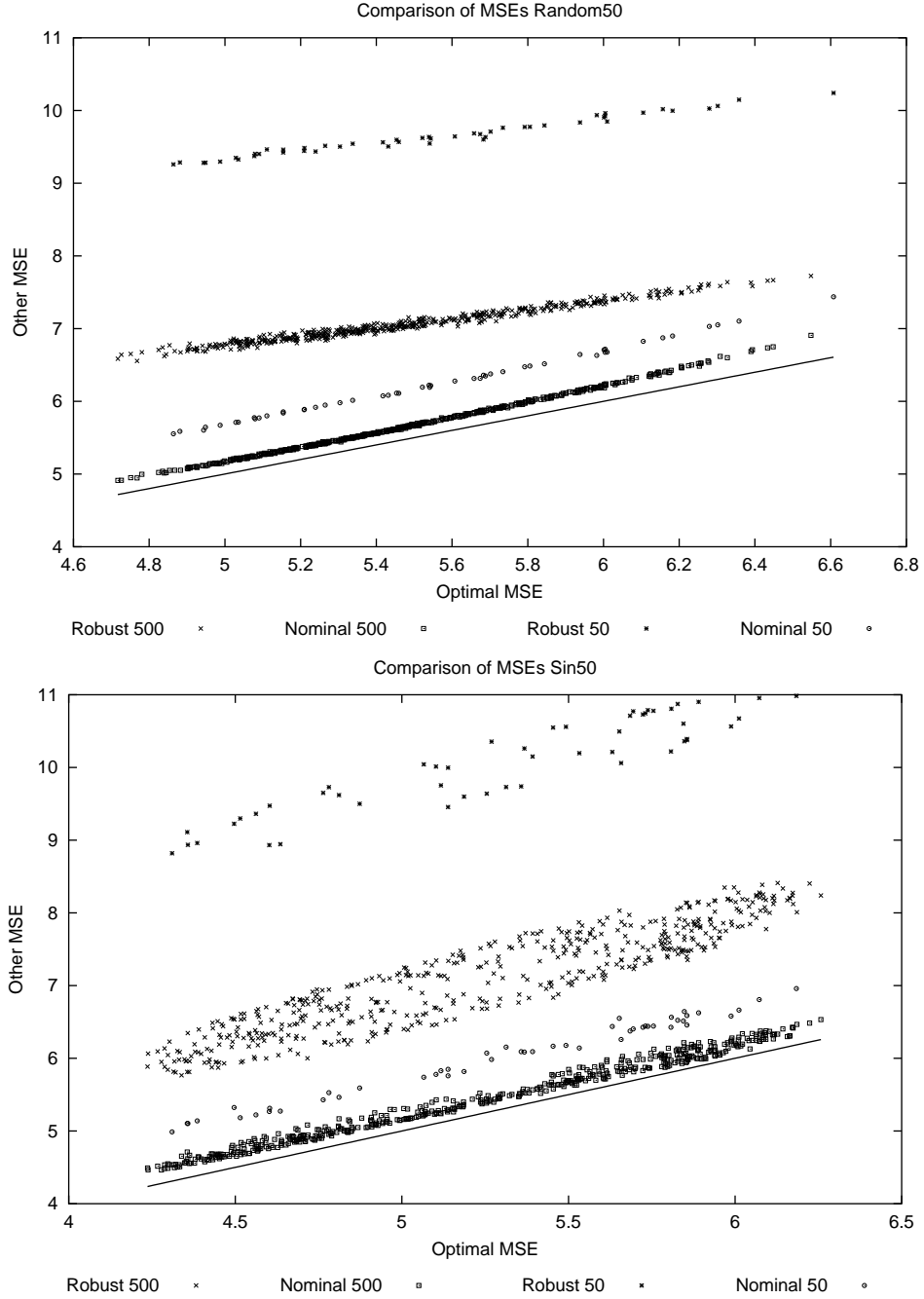


Fig. 5. For a set of either 50 or 500 random PSD pairs from the power-constrained models *Random* and *Sin*, the optimal MSE with Wiener filtering is compared to the MSE for the data-based *Robust* filter and the *Nominal* filter. In all cases the *Nominal* filter has lower MSE than the *Robust* filter. Better spectral estimates and performance are obtained for larger sample sizes.

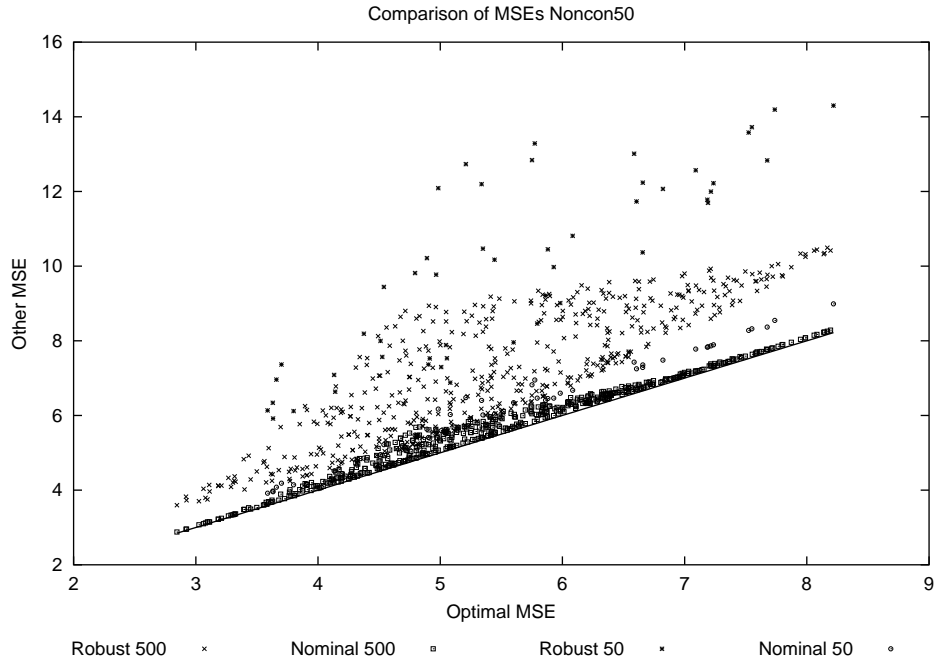


Fig. 6. For a set of either 50 or 500 random PSD pairs from the non-power-constrained model ‘Noncon’, the optimal MSE with Wiener filtering is compared to the MSE for the data-based *Robust* filter and the *Nominal* filter. In most cases the *Nominal* filter has lower MSE than the *Robust* filter. Better spectral estimates and performance are obtained for larger sample sizes.



### C. Summary

Monte Carlo simulations for a three PSD models (two for power-constrained and the specific non-constrained model used for the second robust filter) were conducted. For random PSD pairs from the classes, the simulations compared the optimal MSE achievable to the MSEs of the *Robust* and *Nominal* filters. When the exact model PSDs are known, the *Nominal* filter provides only slightly worse performance than the optimal. The *Robust* filter almost always does worse. This conclusion holds for all three models at SNRs of 5 and 10 and  $b = 0.25$  and 0.5. This result is due to the underlying structure and distributions of the PSD models. Given these models, the performance of the *Nominal* filter does not degrade rapidly enough for deviations from the nominal to make the *Robust* filter useful.

The second set of simulations explored calculating the *Robust* and *Nominal* filters from PSDs as estimated from small data sets. In this case the estimate of the nominal (mean) PSDs is less noisy than that of the lower and upper bounds. Unsurprisingly, as the number of samples in the data set increases, the performance of both robust filters improves. The *Nominal* filter still out-performs the *Robust*, and its performance approaches that of the true *Nominal* filter faster.

These simulations indicate that, assuming any of these models for the PSDs of the signal and noise are reasonable, the best approach to robust filtering is to use the *Nominal* filter derived from the mean signal and noise PSD estimates.

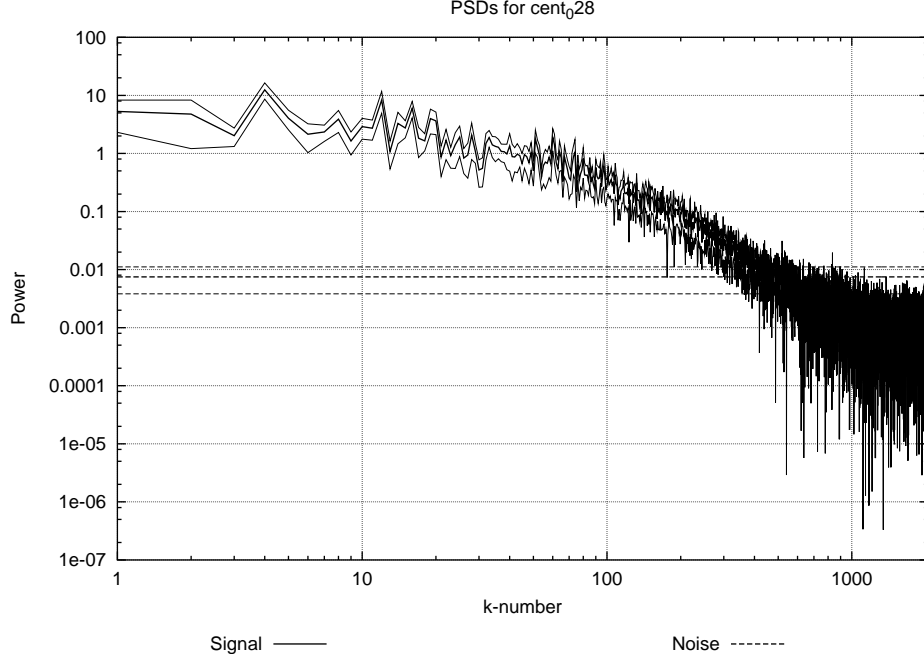


Fig. 7. Estimated PSDs (mean and upper and lower bounds) for experimental data over a set of 28 measurements.

#### 4. Observational Results

Open-loop WFS data for the Lick system was taken on March 10, 2004. Telemetry of slope measurements from 28 subapertures over 4096 consecutive time steps was taken in several trials. These slopes are correlated measurements of the temporally evolving phase. Given these measurements, the *Robust* and *Nominal* filters are estimated from the data. This is a fairly small sample size, so we expect the PSD estimates to be noisy.

The estimated model PSDs are shown in Fig. 7. At very low frequencies the signal PSD does not follow a power law like in our model. The PSD estimates are quite noisy at high frequencies. These results conform to the parameters used in the simulation analysis above in terms of SNR,  $b$  and overall PSD shapes. The filter derived above are both applicable to this model. The ratio of the *Robust* and *Nominal* filters is plotted in Fig. 8. The two filters differ primarily in that the *Robust* filter has slightly lower gain in the mid-range of frequencies for moderate SNRs and much higher gain at the noisiest frequencies. This translates into signal estimates that, in the *Robust* case, follow the high-frequency oscillations of the noisy measurements more closely.

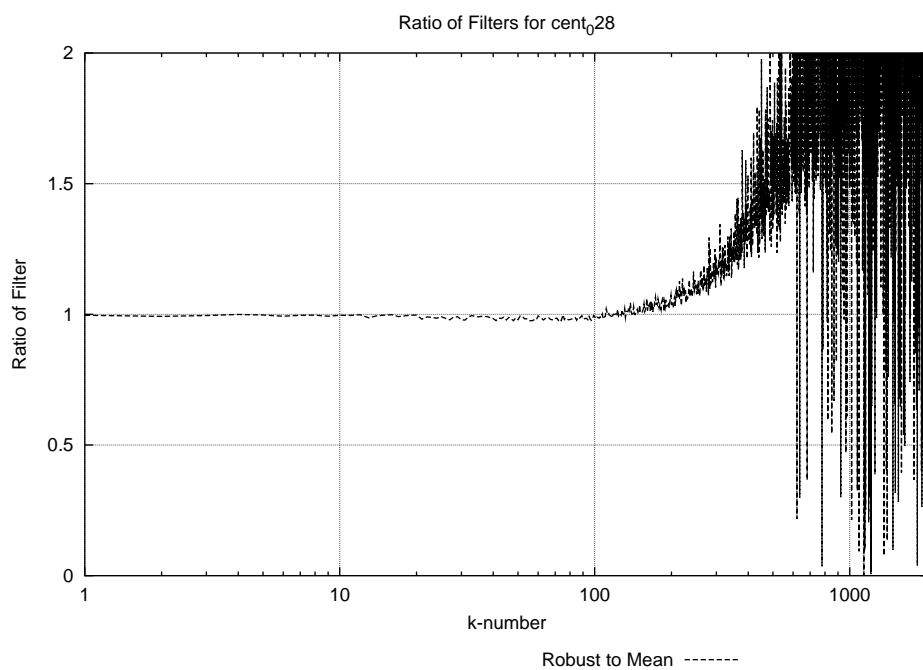


Fig. 8. Ratio of the *Robust* filter to the *Nominal*. At midrange frequencies, the gain of *Robust* is slightly lower. At high frequencies, which are dominated by noise, the gain is much higher.

## 5. Conclusions

The framework of robust filtering in the face of spectral uncertainty has been examined for our specific case of WFS measurements from an Adaptive Optics system. Given two different reasonable models for the classes of signal and noise PSDs, the robust filters were derived. Both models assume that the PSDs are bounded from above and below by finite functions. If the total power of any class member is constrained, the robust filter can be derived using the formulae of Kassam and Lim. The *Robust* filter is pessimistic and lowers the filter gain in high-SNR regions, while increasing the gain in low-SNR regions. If the total power is not constrained, the approach of Poor, with an added assumption about the scaling of the nominal PSD, produces a robust filter which is simply the Wiener filter for the nominal case.

Both these filters were examined in simulation. Monte Carlo simulations show that for our models of class distribution using a power law for the signal PSD and constant noise, the *Nominal* filter performs very well, and almost always outperforms the *Robust* filter. Further Monte Carlo simulation examined the case of estimating the PSDs from a small set of observations. The estimate of the mean PSD is less noisy, and the performance of the *Nominal* filter is better than the *Robust*, especially as sample sizes become small.

Finally, real observations from the Lick AO system were examined. The power spectra for these data are similar to that of the models, differing only at low frequencies by having less power than the power-law model. The robust filters were calculated, and the *Robust* filter differs from the *Nominal* primarily in having much higher gain at high frequencies, where noise dominates, which is not desirable.

Why doesn't the *Robust* filter provide a higher level of performance in our simulations? The construction of the robust filter assumes that for non-nominal PSDs, the performance of the nominally-optimal Wiener filter degrades rapidly. In the three classes that we examined in this project, that was definitely not the case. The *Robust* filter does perform better than the *Nominal* in a few cases, but these rarely occur. This is most likely due to the large differences in signal and noise power in the different bands. In our models (and in our experimental data) the high SNR regions have very high SNR, and the low SNR regions have very low SNR. The pessimism of the robust filter is therefore misguided.

## Acknowledgments

This work was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48. The document number is UCRL-TR-204789. Lisa A. Poyneer can be reached at poyneer1@llnl.gov.

## References

1. NSF Center for Adaptive Optics, <http://cfao.ucolick.org>
2. J. W. Hardy, *Adaptive Optics for Astronomical Telescopes*, (Oxford University Press, New York, 1998).
3. F. Roddier, Ed., *Adaptive Optics in Astronomy*, (Cambridge University Press, Cambridge, 1999).
4. M. H. Hayes, *Statistical Digital Signal Processing and Modeling*, (John Wiley and Sons, Hoboken, New Jersey, 1996).
5. S. A. Kassam and H. V. Poor, "Robust Techniques for Signal Processing: A Survey," *Proc. IEEE* **73**, 433-481 (1985)
6. S. A. Kassam and T. L. Lim, "Robust Wiener Filters," *J. Franklin Institute* **304**, 171-185 (1977)
7. H. V. Poor, "On Robust Wiener Filtering," *IEE Trans. Auto. Control* **25**, 531-536 (1980)